

University of Stuttgart
Germany

Institute of Software Technology
Reliable Software Systems

Modelling and Predicting Memory Behavior in Parallel Systems with Network Links – Palladio-based Experiment Report

Symposium on Software Performance 2019

Philipp
Gruber &
Markus
Frank



Motivation

Lewis
Software Engineer

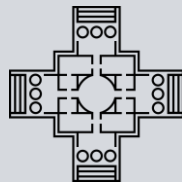


Machine



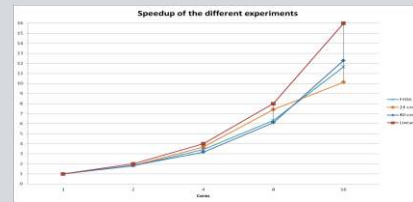
100 Cores

Palladio (PCM)

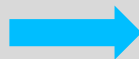
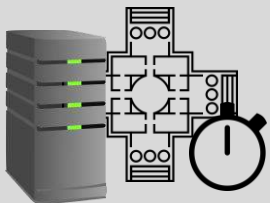


- User
- Hardware
- Software

Performance prediction



Reality:

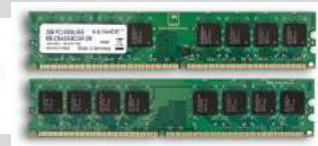
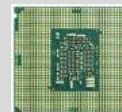


Linear speedup

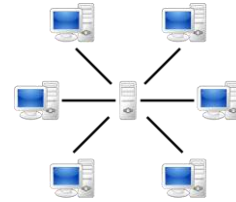
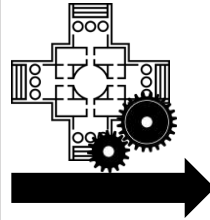
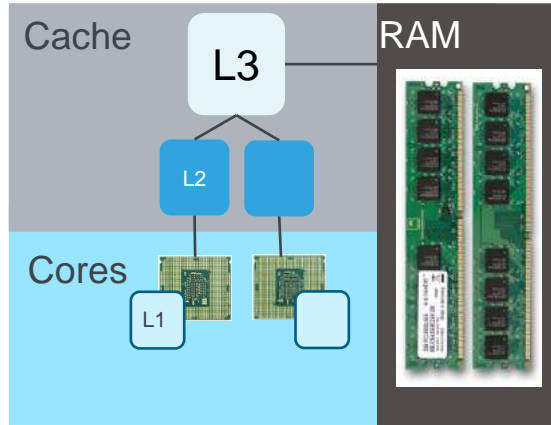


[Frank16]

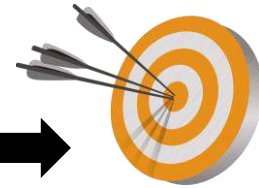
Why?



Approach



Palladio network links
as memory buses



Hypothesis/Goal:
Improve Accuracy

Problem

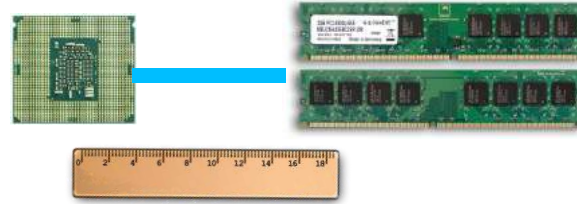
Solution

Result

- RQ1 → Is this possible?
- RQ2 → Is it more accurate

Solution

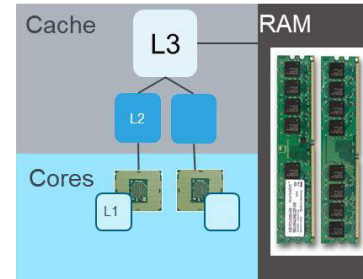
- Modelling the memory bandwidth



How ?

- MemTest86 – RAM Benchmark → Throughput

- Model MemTest in Palladio

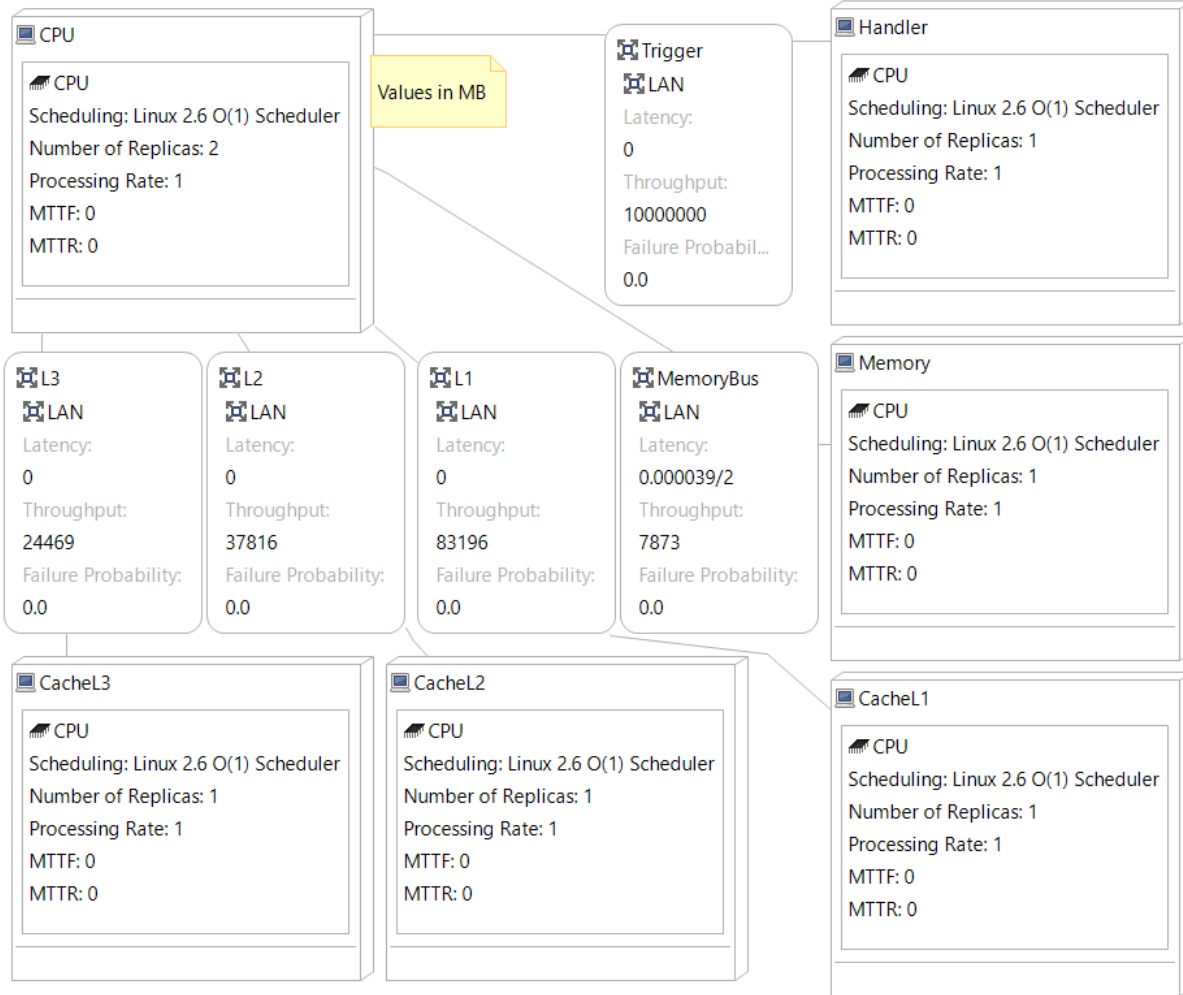


Solution

RQ1 → Is this possible?

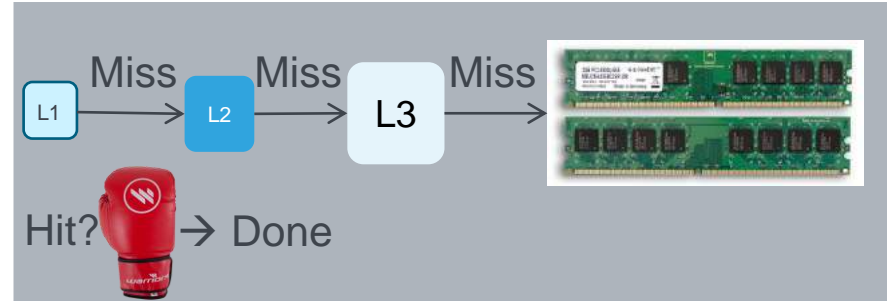


Response time
=
Measurement



Solution

- Modelling the memory behavior





How



- Estimation model of miss/hit ratio

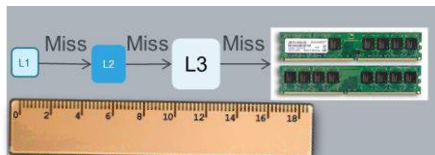


Ratio: 70:30  

Lack in cache behavior information

Solution

- Measuring How ?



- Repetition of [Frank16]'s experiment



small & big



machine

- Perf – Linux tool



L1 & LLC(L3) loads



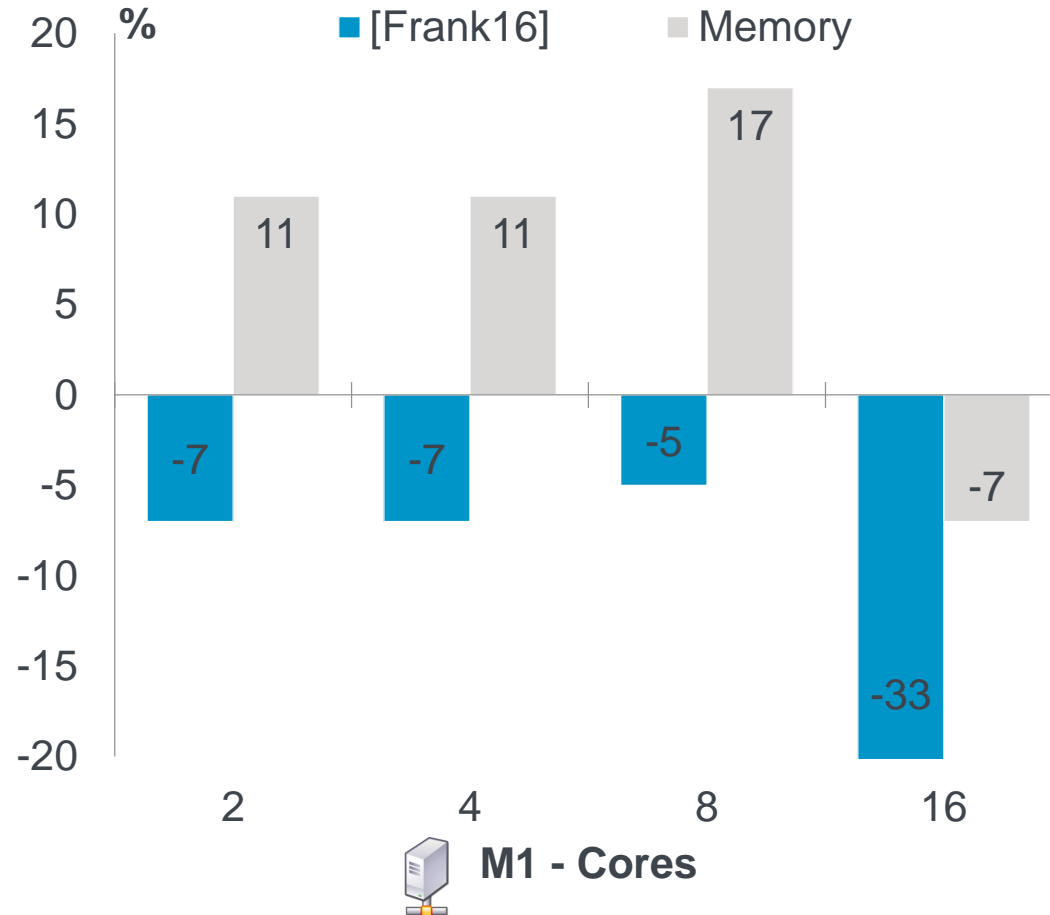
	1 Core	16 Core
Small (L3)	33:67	55:45
Small (L1)	75:25	70:30

- Plug in Palladio (repository model)

Results

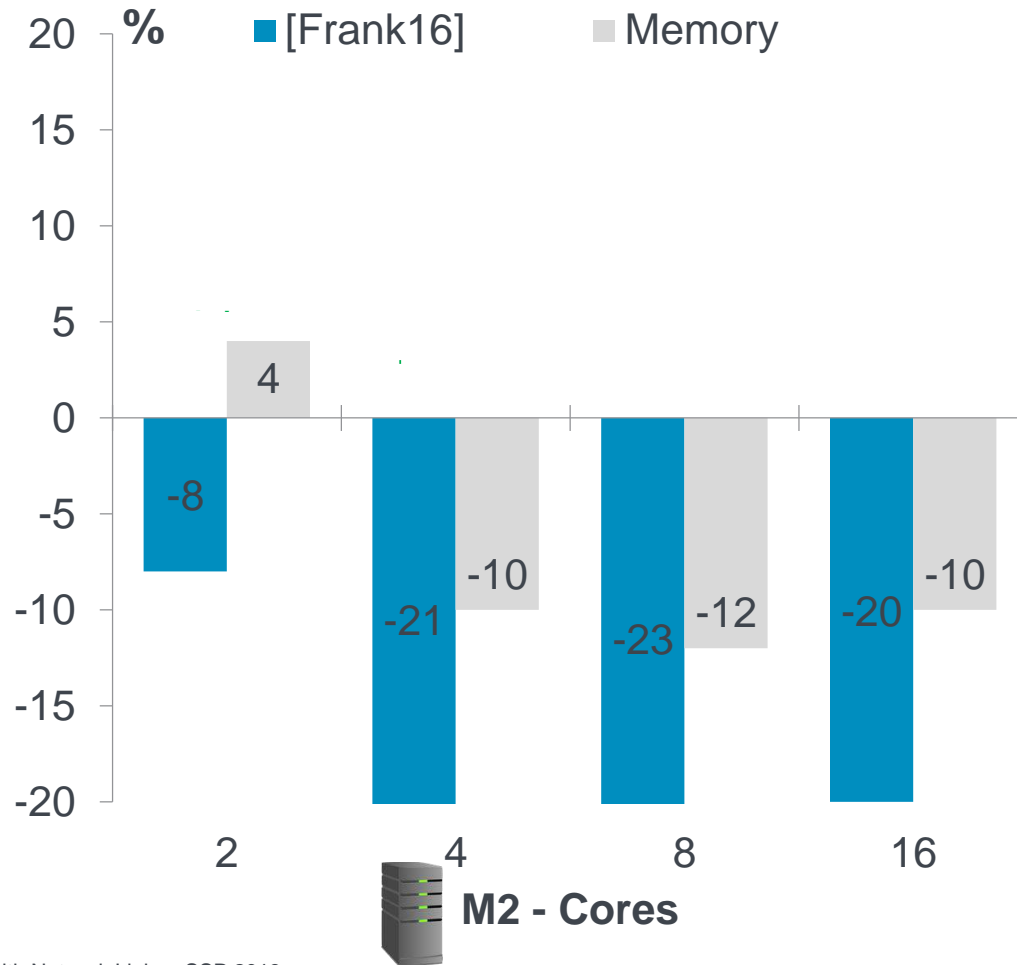
Results

Prediction error



Results

Prediction error

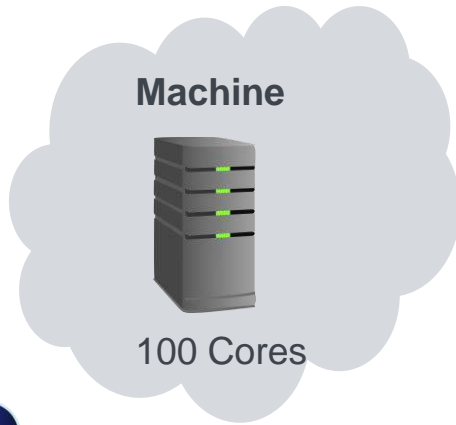


Summary

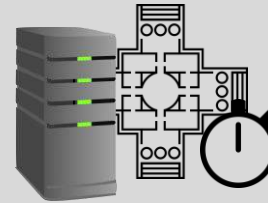
Lewis
Software Engineer



Task: Ensure
Quality attributes



Reality:



Linear
speedup



[Frank16]

Lessons learned:

- Proof of concept
- Accuracy gain
- Remaining inaccuracy



Future work:

- Hardware behavior
- Reduce the calibration effort
- Meta model changes

Back up

Result

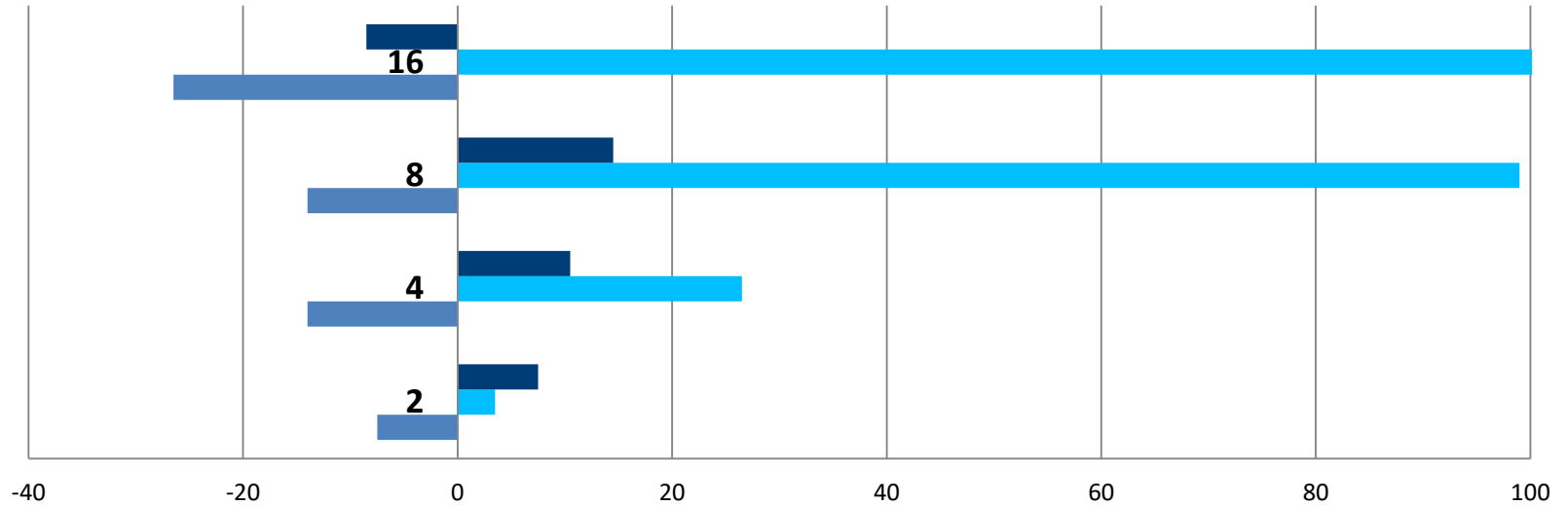
Measured cache ratios

	1 Core	2 Core	4 Core	8 Core	16 Core
Small (L3)	33:67	34:66	54:36	80:20	55:45
Small (L1)	75:25	75:25	75:25	75:25	70:30
Big (L3)	24:76	20:80	34:66	77:33	85:15
Big (L1)	70:30	70:30	70:30	70:30	70:30

Result

Prediction error

Cores



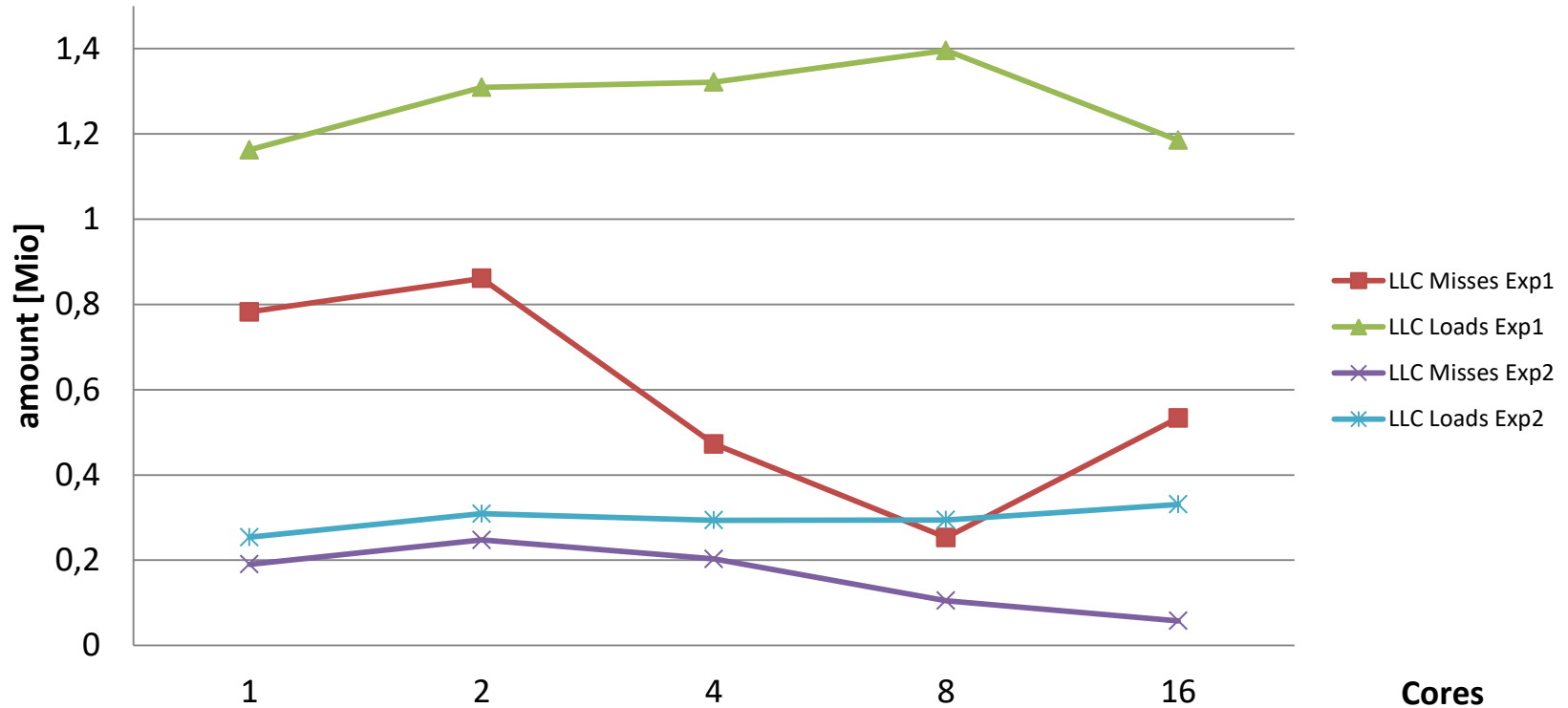
- 10% mean prediction error, before 15 %

→But ...

Cache measuring

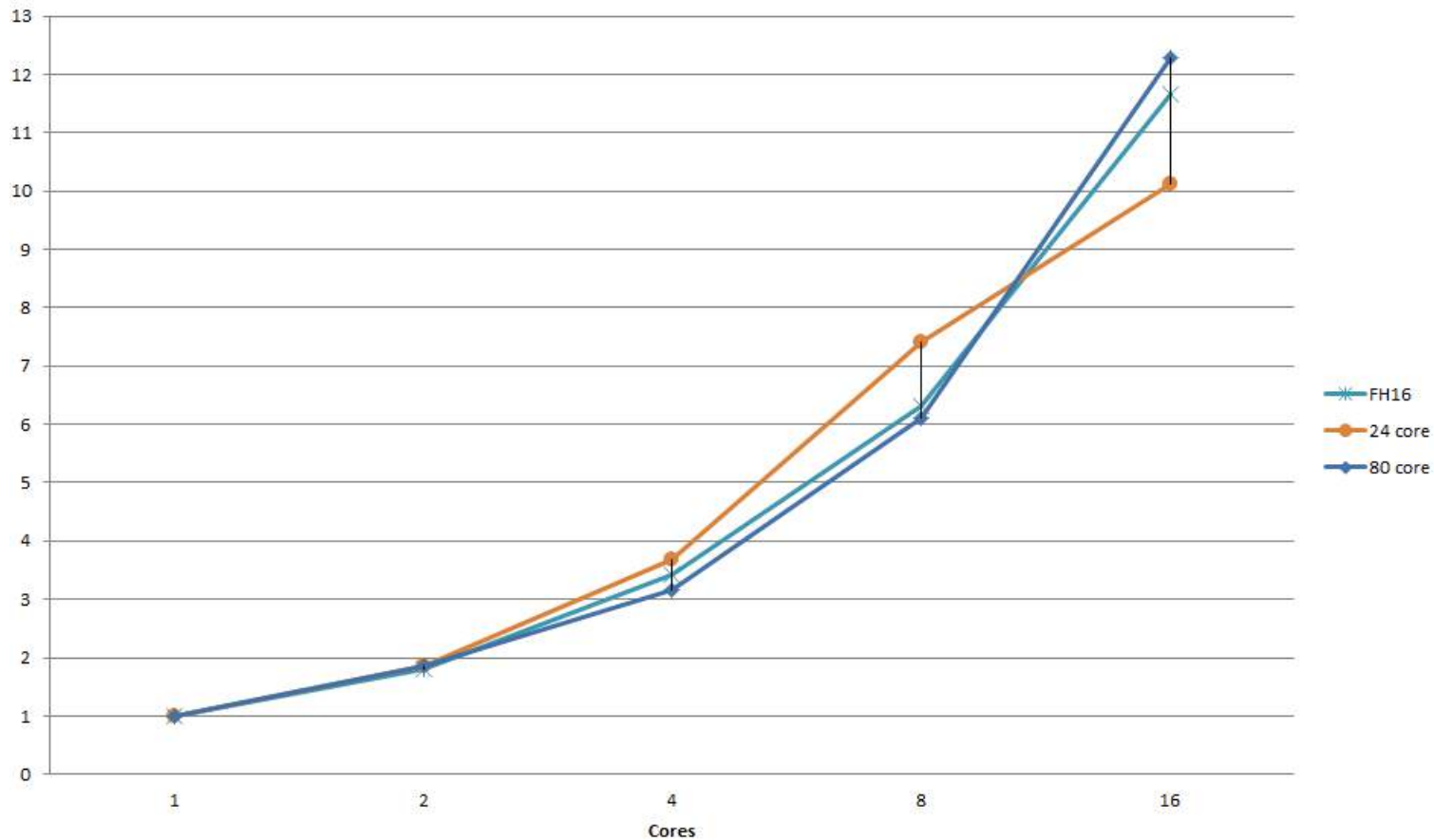
Hardware „Potsdam“: 2*6 (2.5 GHz) vs. 4*10 (2.4 GHz)

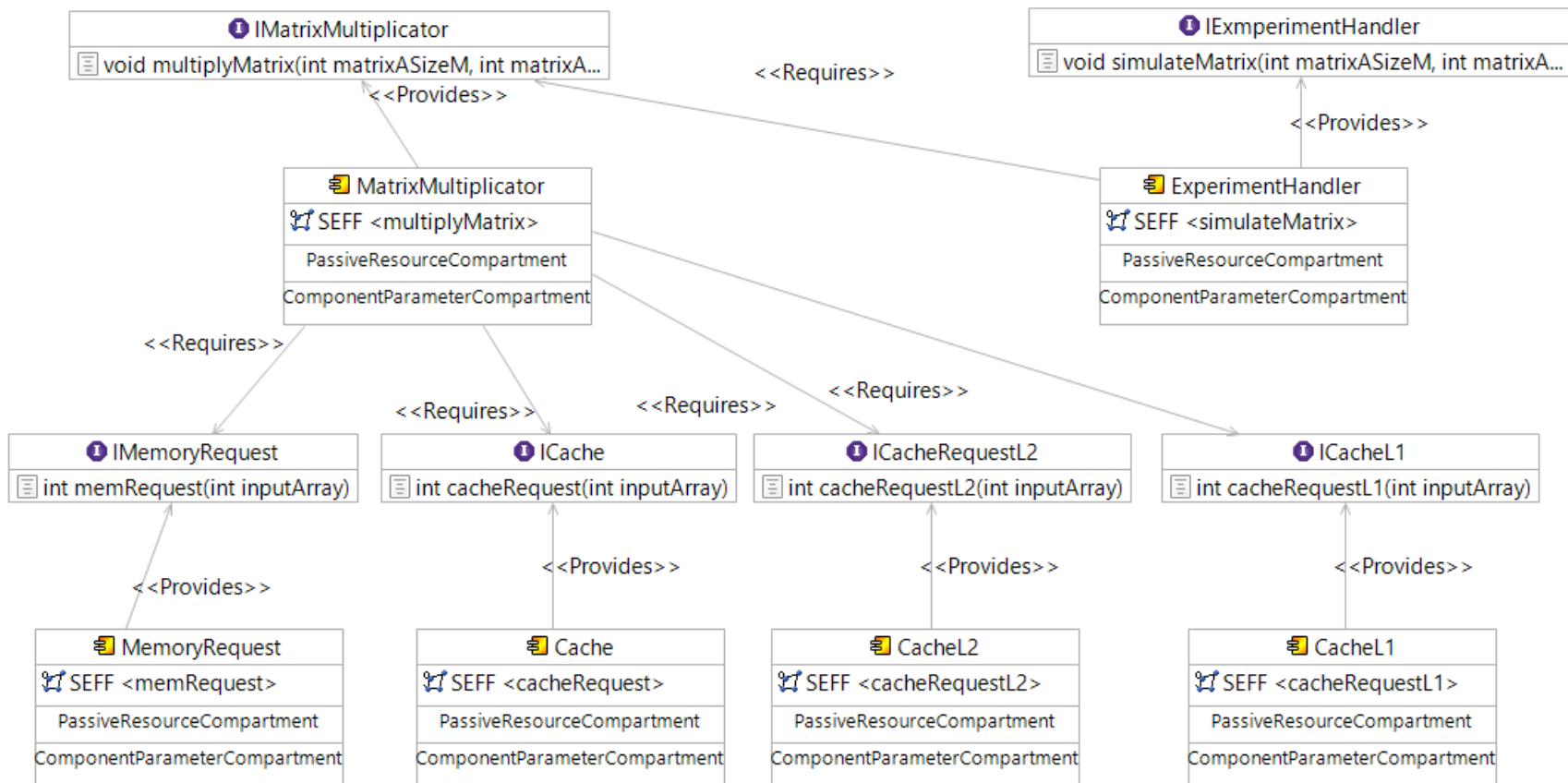
- Cache: L1: 32 KB, L2: 256 KB, L3: 15 MB vs. 30 MB
- RAM: 24 GB vs. 896 GB



- Factors: Amount of modules, L3 Size, Clock frequency

Speedup of the different experiments





Copyright

- „Villa Rotonda“ from <https://commons.wikimedia.org/wiki/File:01-Villa-Rotonda-Palladio.jpg>
- Cliparts from <http://office.microsoft.com/de-de/images/>
- Palladio logo from <https://sdqweb.ipd.kit.edu/wiki/File:Pcm-logo-stilisiert.png>
- Rest made by myself

References

[Frank16] – M. Frank and M. Hilbrich „Performance Prediction for Multicore Environments – An Experiment Report”, 2016